



Ù&@å~|ā * Á@ÁÓ^|* ãð Á[&&!Á^æ~^
Á

ŸÖ!ā•ÁÕ[[••^}•Á}Áā•Á] ā\•{ æ

DEPARTMENT OF DECISION SCIENCES AND INFORMATION MANAGEMENT (KBI)

Scheduling the Belgian Soccer League

D.R. Goossens*, F.C.R. Spieksma

ORSTAT, K.U.Leuven, Naamsestraat 69, 3000 Leuven, Belgium

Abstract

Any sports league needs a schedule of play, and such a schedule can be important, as it may influence the outcome of the sports competition itself and since it may have an impact on the interests of all parties involved. As in many other sports leagues and countries, the interest for Belgian soccer has increased over the last years. This paper describes our experiences in scheduling the highest Belgian soccer league. We describe how we automated and improved the way in which the calendar is constructed, resulting in the schedule for the season 2006–2007. We also explain how decomposing the scheduling problem into two subproblems resulted in a further improvement. The resulting calendar has been accepted for the season 2007–2008. Finally, we compare the quality of the schedules resulting from the different methods.

Keywords: scheduling, sports, soccer, integer programming

1 Introduction

Especially in Europe, soccer has become big business, involving many stakeholders (e.g. teams, police, fans, broadcasting companies, etc.) and a lot of money. In Belgium, Belgacom TV pays a historic 12 million per year for the soccer broadcasting rights. This is a trifle compared to the amounts that are paid for broadcasting the soccer competition of countries like e.g. the United Kingdom (£900 million) or compared to other sports like baseball (>\$500 million) and basketball (>\$600 million) in the United States. Nevertheless, it illustrates a rising interest in Belgian soccer and with that, an increased importance of the league schedule. Indeed, apart from the obvious influence on the results of the sports competition itself, the schedule also affects the attendance and public interest for the competition, and with that the profitability of the event for broadcasters, sponsors, and advertisers. Each involved party has its (possibly conflicting) constraints and wishes, which makes it challenging to come up with a schedule that is considered fair and acceptable to all parties. In literature, there are papers on the scheduling of the national soccer league of various countries, e.g. Germany and Austria (Bartsch, Drexl & Kroger 2006), Italy (Della Croce & Oliveri 2006), The Netherlands (Schreuder 1992), Denmark (Rasmussen 2006), Chile (Noronha, Ribeiro, Duran, Souyris & Weintraub 2007), and Brazil (Ribeiro & Urrutia 2006). However, because of some specific constraints that characterize each of these competitions, the models presented in

*Corresponding author. E-mail: Dries.Goossens@econ.kuleuven.be

these papers are not readily applicable to soccer league scheduling problems in other countries.

2 Problem description

The Jupiler League is the highest division in Belgian soccer and consists of 18 teams. The league is intermural, meaning that a team plays a game in its own stadium or in the stadium of its opponent. The teams play a double round robin tournament, i.e. each team plays against each other team twice, once at home and once away. Furthermore, teams should not play more than two consecutive home (away) matches (i.e. breaks), and the total number of breaks should be minimal. No team should start or end the league with a break. For reasons of fairness, the schedule in the second half of the competition should be the same as in the first half, but with the home advantage inverted (i.e., mirroring). The league is temporarily constrained, meaning that the number of matchdays (or rounds) on which a game may be played equals the number of games each team has to play. Moreover, the 34 weekends on which there is a matchday are given by the calendar committee. This committee is installed by the Royal Belgian Football Association (KBVB), and consists of representatives of 6 of the 18 clubs. The calendar committee is responsible for producing a schedule 6 weeks before the start of the competition.

Several requirements, originating from three stakeholders, should be taken into account. These requirements are based on conversations we had with Robert Sterckx (KBVB), secretary of the calendar committee. First of all, a mayor can forbid that a game is played in his or her city on one or more dates. The reason behind this is usually that there is some other event (e.g. a fair or a pop concert) in the city needing the attention of the local police. A number of games, called *risk games*, are known to have an increased risk of hooliganism. Since the additional police forces needed to guarantee the safety of a risk game may not be available on some date, a mayor may forbid a risk game, but allow a regular game to take place on that date.

Second, clubs have a wide variety of wishes. For instance, some teams prefer not to play at home when some other team plays at home, because they fear that a part of their spectators would attend the other game, or simply because they share a stadium. Teams also have a number of wishes related to the fairness of the schedule (e.g. no team wishes to face all the traditionally strong opponents in a row), or wishes related to the European schedule (e.g. some clubs prefer a home game after a game for a European competition), or wishes related to the expected number of spectators (e.g. no top game in the summer, when many fans are abroad for holidays) or they simply want to increase their chances of playing a good season (e.g. most clubs prefer to start the season with a home game and easy opponents).

Third, the TV station that acquired the broadcasting rights (Belgacom TV) also has its demands, intended to maximize the viewing figures. For instance, broadcasters wish to have the top games (i.e., games between two of the four *big*

clubs in Belgium) spread equally over the season and prefer a competition that is thrilling until the final whistle. Matches are normally played on Saturday, however, the broadcasting company has the right to shift a match to Friday and two matches to Sunday on a month's notice. Since this could be problematic for teams that also play a midweek match (e.g. Champions League), matches between those teams should be scheduled as much as possible on weekends that are not preceded or followed by a midweek match.

The scheduling problem is to decide for each matchday which teams play against each other and which one of each pair plays at home, satisfying as many requirements as possible.

3 Scheduling the soccer league by hand

In the traditional approach (used up to the season 2005–2006) a schedule was constructed by hand, under supervision of the secretary of the calendar committee. The starting point of this manual approach is a so-called basic match schedule (BMS). A BMS gives for each team a home-away assignment and the opponent. Table 1 shows the first 7 rounds of the BMS that has been used for decades in Belgian soccer. The schedule shows that team 1 opens the season with a home game against team 3. On the second matchday team 1 plays an away game against team 5, and on the third matchday team 1 receives team 7. This basic match schedule is a so-called canonical schedule (De Werra 1980), which satisfies the mirroring constraint and minimizes the number of breaks.

1	2	3	4	5	6	7
1-3	2-4	1-7	2-8	1-11	2-12	1-15
4-17	3-18	3-5	4-6	3-9	4-10	3-13
6-15	5-1	6-2	5-18	5-7	6-8	5-11
8-13	7-16	8-17	7-3	8-4	7-18	7-9
10-11	9-14	10-15	9-1	10-2	9-5	10-6
12-9	11-12	12-13	11-16	12-17	11-3	12-4
14-7	13-10	14-11	13-14	14-15	13-1	14-2
16-5	15-8	16-9	15-12	16-13	15-16	16-17
18-2	17-6	18-4	17-10	18-6	17-14	18-8

Table 1: Basic match schedule (matchdays 1-7)

Clearly, when all teams have been assigned a number in the BMS, a schedule follows. This task is carried out by hand, using merely an Excel sheet to visualize the assignments, and lots of patience. First, a number from the BMS is assigned to the each of the four top clubs, bearing in mind a reasonable spread of their mutual games, and the police constraints. Next, one by one a number is assigned to each of the other teams, again trying to satisfy the police demands. Afterwards, an attempt is made to improve the schedule by swapping (manually) the assignments of a pair of teams.

Not surprisingly, a calendar resulting from this approach satisfied only a minority of the constraints, and especially the wishes from the clubs and the TV station were often brushed aside. As a result, a number of teams have expressed their displeasure with the league schedule, calling it unbalanced and unfair (see e.g. Lambaerts (2005)). Moreover, it was unclear, even to teams that had a representative in the calendar committee, what wishes were taken into account or how the eventual schedule was constructed. This even led to insinuations that the chairman of the calendar committee was favouring his own team (see Reunes (2005)).

4 Automating the scheduling procedure

After it was illustrated in a Master's thesis (Demasure 2006) that there was ample room for improvement of the schedule for the then ongoing season 2005–2006, we were invited by the KBVB to work on the schedule for the new season.

First of all, we sought to clarify some of the constraints in consultation with the calendar committee and the involved parties. The TV station expressed its wish for a thrilling season in concrete by asking that on each matchday, at least one (and preferably two) of the top teams would play an away game. The underlying motivation is that a top team's home games are less interesting with respect to viewing figures, since the top team tends to win these games without much tension. As an extension of this wish, the TV station prefers that on each matchday, at least 3 out of the 6 teams that finished on top of the league the previous season (which usually include the four top teams) play an away game. Furthermore, they wish to have only one top game per matchday, and no top games in the first four matchdays. Next, the great variety of wishes and complaints by the clubs needed to be curtailed. After some consideration, the clubs limited their constraints to the following four types:

- no home/away game on a given matchday m
- no home game when some other team x plays at home
- home game against team x in the first/second half of the season
- no home game against a top team on a given matchday m

Additionally, the calendar committee came up with a number of constraints to be taken into account. Since most teams prefer not to play against all top teams consecutively, the committee asked that no team would have to play more than twice against a top team in four consecutive games. Furthermore, the calendar committee stated that every team should receive a top team at home at least once in each half of the season. In that way, the calendar committee gave in to the main objection by the clubs (Lambaerts 2005). Finally, in order to protect the teams that also play midweek games in one of the European competitions, another constraint was added stating that these teams should not play a top game on a matchday immediately before a European game.

Second, as the wishes and constraints kept coming in, it turned out that satisfying them all was simply not possible since they were conflicting. Therefore, we asked the calendar committee to attach one of five priority levels to each of the wishes. The highest level was reserved for constraints arising from physical impossibilities (e.g. two teams sharing the same stadium, or a stadium not being available due to maintenance works). The second highest level was used for the wishes coming from the police and local governments, since after all, a mayor in Belgium has the right to forbid a game if he or she thinks the safety cannot be guaranteed. The calendar committee made an assessment between the wishes of the clubs and the TV station, balancing the underlying financial or sporting motives with the fairness of the schedule, and assigned them to one of the three remaining priority levels.

Finally, we constructed a mixed integer programming model, automating the assignment of numbers in the basic match schedule to teams as was earlier done by hand. We linked a penalty with each priority level, such that this penalty was incurred if the corresponding constraint was violated (we refer to the appendix A for a more detailed discussion of this model). The optimal schedule then was the schedule following from the assignment with minimal incurred penalties. After a number of meetings with the calendar committee and representatives of the broadcasting company, we settled upon a choice for the penalties and the priority levels, and we came to a result with which the calendar committee was satisfied. The resulting calendar was officially presented to the press as the calendar for the season 2006–2007.

5 Beyond the basic match schedule

The basic match schedule has been used by the Royal Belgian Football Association (KBVB) for decades, to schedule not only the first division, but also the 7 other national divisions and numerous regional leagues. Furthermore, the schedules for these leagues are not independent of the first division schedule, since many lower division teams prefer not to play a home game when a neighboring higher division team plays at home. Since the first division is by far the most constrained league, it is scheduled first. Afterwards, the second division is scheduled, and so on, following the hierarchy among the divisions. Since people involved in scheduling these leagues are highly familiar with the BMS, we were asked to stick to it for the 2006–2007 season schedule.

Despite the fact that the basic match schedule has many interesting properties (e.g. a minimal number of breaks), it is obviously also highly restrictive. After careful consideration with the lower division schedulers, we found out that the wish of lower division teams not to play at home at the same time as a neighboring higher division team was in fact the only relevant dependency between the leagues. This concern however does not require using the basic match schedule, since it can also be accomplished by using the same home–away patterns as in the basic match schedule. These home–away patterns state for each round whether a team plays a home or an away game, and for each home–away pattern in the BMS, there is another home–away pattern in the BMS that is its exact

opposite.

Using the home-away patterns as a starting point allows us to use a 2-phase approach, where in the first phase each team is assigned a home-away pattern and in the second phase, the actual opponents are decided. In phase 1, only constraints that relate solely to whether a team plays at home or away can be considered. These constraints are the complementarity constraints (i.e., two teams do not want to have home games at the same time), place constraints (i.e., a team does not want a home/away game on a given matchday), and the TV constraints requiring for each matchday a minimal number of top clubs playing an away game. The first phase results in a home-away pattern for each team, which is used as an input for the second phase. In this phase, all other constraints, i.e. those that depend on actual opponents, are taken into account. For instance, enforcing that some team does not play against a top team on a given matchday can be done in this phase. The result of this phase is the actual schedule. We used the 2-phase approach to construct the schedule for the season 2007–2008 and refer to Appendix B for mixed integer programs for both phases.

Sports scheduling problems are often decomposed into multiple smaller sub-problems, which are then solved sequentially. In general, we can distinguish between “first-schedule-then-break” approaches, where first the opponents for each team are determined and afterwards it is decided who plays at home, and “first-break-then-schedule” approaches, where the home-away pattern of each team is decided first and the actual opponents are determined afterwards. Both methods have been used in literature (see Rasmussen & Trick (2007) and Briskorn (2007) for an overview), but especially for scheduling Belgian soccer, a first-break-then-schedule method seems quite appropriate. Indeed, the constraints that are taken into account while assigning the home-away patterns are in general the most important constraints, since they often originate from the police or from stadium unavailabilities.

One advantage of a phased approach is that the resulting schedule seems more robust to changes in the constraints or priorities than a schedule based on the BMS. Indeed, with the BMS approach, even the smallest change in the input typically results in a totally different schedule, whereas with the 2-phase approach an additional constraint does not necessarily turn the previous schedule upside down. Obviously, this makes the calendar committee’s task less confusing and allows its members to suggest minor adjustments of the schedule. Another issue is that the 2-phase approach schedule tends to suffer less from the so-called carry-over effect than the BMS based schedule. In high-contact sports as soccer, one can imagine that if some team A is very strong, its opponent B will be exhausted after the game against A . Since this team B will play against another team C on the next matchday, we can say that A ’s game against B has an impact on the game against C . In other words, there is a carry-over effect from A to C . Ideally, team A should carry over to team C at most twice during the season, such that the carry-over effects are maximally balanced over the teams. The carry-over effect has been suggested as a reason why one team relegated in the Belgian media (see Geril (2007)), and thus the fact that the 2-phase approach reduces this effect can be considered an advantage.

6 Results

A comparison between the three discussed methods is not straightforward, since they were applied on three different instances of the Belgian soccer scheduling problem. Nevertheless, although the number of constraints in these instances increases over the years, the number of teams and the type of constraints does not. Table retable2 summarizes the most important results.

	Manually (2005–2006)	Assignment (2006–2007)	2-phase approach (2007–2008)
Computation time	1 week	4–5 hours	5–10 minutes
Goal function value	> 75,000	11,698	2,144
Police constraints	70%	95%	100%
Club constraints	32%	68%	66%
Number of rounds with away games for teams in top 4 (2/1/0)	10/24/0	26/8/0	28/6/0
top 6 (3/2/1/0)	2/20/12/0	18/16/0/0	20/14/0/0
Teams with top games in both season halves	13	18	18

Table 2: Results of the three solution methods

The assignment model, and especially the 2-phase approach, results in obtaining a schedule quite a lot faster than in the previous manual approach. This is very important, since it gives the calendar committee extra time to suggest improvements and hence allows for alternative schedules to be investigated. Indeed, until 2005, it was a hard enough task to come up with a schedule at all, let alone redoing everything to answer to the concerns of members of the calendar committee. Moreover, there are only a couple of days between the day on which it is known which teams promote to the first division and the day on which the new schedule is to be made public. The computational efficiency of our methods allows creating a schedule for every possible scenario in advance. Once it is clear which teams go up to the first division, the calendar committee just needs to pick the schedule of the corresponding scenario.

According to the goal function value, the quality of the schedules created with the assignment approach is about 7 times higher than the hand-made schedule. The 2-phase approach further improves on the assignment model by a factor of 5. This improvement in quality also shows when we look at the number of constraints that are satisfied. Whereas the schedule made by hand satisfied 70% of the police constraints for the season 2005–2006, the assignment approach increased this to 95% for the season 2006–2007, and using the 2-phase approach we reached 100% for the season 2007–2008. Despite an ever increasing number of club wishes, the proportion that could be satisfied doubled to over two thirds

since scheduling by hand was abandoned. A careful selection of the priorities made it possible to grant at least one wish from each team. In the hand-made schedule, the consideration of TV wishes was limited to a reasonable spread of the top games and the broadcasters were granted no more than 10 matchdays with two top teams playing an away game. Furthermore, they had to cope with 20 matchdays with only two teams out of last year's top 6 playing an away game, and even 12 matchdays with only one of these teams playing away. In the schedule made with the assignment approach, the TV station was offered 26 matchdays with two top teams playing an away game. If we consider last year's top 6, there were no matchdays with only one team playing away, and the number of matchdays with three teams playing away rose from 2 to 18. These results were improved even further with the two-phase approach. This approach also allowed the calendar committee to enforce protection for the clubs playing European midweek games and to avoid successive games against top teams. Finally, all 18 teams now receive at least one top club at home in both halves of the season, whereas with the hand-made schedule, this was the case for only 13 teams.

7 Concluding remarks

The results show that all involved parties have benefited from our approach. Indeed, the police, the clubs and the TV station have been presented a schedule that meets their wishes more than ever before. As a consequence, they gradually come up with more wishes, which keeps scheduling the Belgian soccer a challenging task year after year. Obviously, there remain a number of wishes that cannot be satisfied. However, to those stakeholders that are confronted with an unfulfilled wish, it should be clear that their sacrifice was made in the public interest. Indeed, the computational efficiency of our approach allows us to show the negative consequences with respect to the other constraints if their wish was granted, which makes it easier for them to accept the schedule.

Our approach was received positively in the media. Most newspapers noticed that more clubs have a better schedule and praised the calendar committee for professionalizing the scheduling process (see e.g. Cuvelier (2006)). Nevertheless, negative comments seem unavoidable, as it is in the best interest of the coaches to hedge their team against a poor season start by referring to a touch schedule. Guy Mangelschots, former coach of Sint-Truiden, summarized it as follows: "When after five games you look at the points you have, only then you can say whether the schedule was good or bad" (Martens 2007).

Appendix A: The assignment model

In the mathematical model based on the basic match schedule that we used to schedule the 2006–2007 season, we used the following notation:

$c_{t,n}$ = penalty associated with assigning number n to team t

q_c = penalty associated with violating constraint c

$x_{t,n}$ = 1 if team t is assigned to number n , 0 otherwise

y_c = 1 if constraint c is violated, 0 otherwise

Since not all penalties can be expressed in terms of the x variables (e.g. two teams preferring not to play against each other on a given matchday), we introduced the decision variables y , resulting in the following model:

$$\text{minimize } \sum_t \sum_n c_{t,n} x_{t,n} + \sum_c q_c y_c$$

subject to

$$\begin{aligned} \sum_t x_{t,n} &= 1 & \forall n \\ \sum_n x_{t,n} &= 1 & \forall t \\ \sum_n x_{t,n} &\leq 1 + y_c & \forall c \\ x_{t,n} &\in \{0, 1\} & \forall t, n \\ y_c &\in \{0, 1\} & \forall c \end{aligned}$$

The goal function minimizes the total penalty of the violated constraints. The first set of constraints ensures that each number is assigned to exactly 1 team, while the second set of constraints makes sure that each team is assigned to exactly 1 number. The third set of constraints can be used to model any of the constraints relevant to scheduling the Belgian soccer.

Appendix B: The 2-phase approach

In the first phase, we assign a home-away pattern to each team, using the following notation:

$c_{t,p}$ = penalty associated with assigning home-away pattern p to team t

q_c = penalty associated with violating constraint c

$x_{t,p}$ = 1 if team t is assigned to home-away pattern p , 0 otherwise

y_c = 1 if constraint c is violated, 0 otherwise

Apart from the definition of the x -variables, the resulting model looks very similar to the assignment model, however, the number of constraints that is taken

into account in this phase is considerably smaller.

$$\text{minimize } \sum_t \sum_p c_{t,p} x_{t,p} + \sum_c q_c y_c$$

subject to

$$\begin{aligned} \sum_t x_{t,p} &= 1 & \forall p \\ \sum_p x_{t,p} &= 1 & \forall t \\ \sum_p x_{t,p} &\leq 1 + y_c & \forall c \\ x_{t,p} &\in \{0, 1\} & \forall t, p \\ y_c &\in \{0, 1\} & \forall c \end{aligned}$$

The goal function minimizes the total penalty of the violated constraints. The first set of constraints ensures that each home-away pattern is assigned to exactly 1 team, while the second set of constraints enforces that each team is assigned to precisely 1 home-away pattern. The third set of constraints can be used to model any of the constraints relevant in this phase.

In the second phase, the actual games are determined, using the home-away pattern assignment from phase 1 as an input. We use the following notation:

$c_{i,j,k}$ = penalty associated with team i playing at home vs. team j on round k
 q_c = penalty associated with violating constraint c
 $x_{i,j,k} = 1$ if team i plays a home game vs. team j on round k , 0 otherwise
 $y_c = 1$ if constraint c is violated, 0 otherwise

The model for the second phase decides on the actual opponents of each team on every matchday. We point out that an $x_{i,j,k}$ variable only exists insofar a home game of team i against team j on round k is allowed by the home-away patterns that were assigned to teams i and j . Notice also that because of the symmetry assumption, we only need to schedule the first half of the season.

$$\text{minimize } \sum_i \sum_j \sum_k c_{i,j,k} x_{i,j,k} + \sum_c q_c y_c$$

subject to

$$\begin{aligned}
\sum_j (x_{i,j,k} + x_{j,i,k}) &= 1 & \forall i, k \\
\sum_k (x_{i,j,k} + x_{j,i,k}) &= 1 & \forall i, j : i \neq j \\
\sum x_{i,j,k} &\leq 1 + y_c & \forall c \\
x_{i,j,k} &\in \{0, 1\} & \forall i, j, k \text{ allowed by the HA patterns of team } i \text{ and } j \\
x_{i,j,k} &= 0 & \forall i, j, k \text{ not allowed by the HA patterns of team } i \text{ and } j \\
y_c &\in \{0, 1\} & \forall c
\end{aligned}$$

The goal function minimizes the total penalty of the violated constraints. The first set of constraints makes sure that each team plays exactly once on each matchday, while the second set of constraints ensures that each pair of teams meet each other precisely once in each half of the season. Again, the third set of constraints can be used to model any of the constraints relevant in this phase.

References

- Bartsch, T., Drexler, A. & Kroger, S. (2006). Scheduling the professional soccer leagues of Austria and Germany, *Computers and Operations Research* **33**(7): 1907–1937.
- Briskorn, D. (2007). *Sport leagues scheduling: models, combinatorial properties, and optimization algorithms*, PhD thesis, Christian-Albrechts-Universitt Kiel.
- Cuvelier, M. (2006). Voetbalkalender voor volgend seizoen bekend, Het Nieuwsblad, June 12th (VUM).
- De Werra, D. (1980). Geography, games and graphs, *Discrete Applied Mathematics* **2**(4): 327–337.
- Della Croce, F. & Oliveri, D. (2006). Scheduling the Italian Football League: an ILP-based approach, *Computers and Operations Research* **33**(7): 1963–1974.
- Demasure, S. (2006). *Wedstrijdplanning van de nationale voetbalcompetitie*, Master’s thesis, K.U.Leuven, Faculty of Economics and Applied Economics.
- Geril, J. (2007). Ons budget voor transfers? Nul komma nul euro, Het Nieuwsblad, February 2nd (VUM).
- Lambaerts, G. (2005). Voetbalkalender eerste klasse 2005-2006: Genk boos, De Standaard, June 15th (VUM).
- Martens, L. (2007). Wij zijn niet ontevreden, Het belang van Limburg, June 20th (Concentra).

- Noronha, T., Ribeiro, C., Duran, G., Souyris, S. & Weintraub, A. (2007). A branch-and-cut algorithm for scheduling the highly-constrained Chilean soccer tournament, *Lecture Notes in Computer Science*, *in press* .
- Rasmussen, R. (2006). Scheduling a triple round robin tournament for the best Danish soccer league, *Technical report 2006/1*, University of Aarhus, Department of OR.
- Rasmussen, R. & Trick, M. (2007). Round robin scheduling – a survey, *European Journal of Operational Research*, *in press* .
- Reunes, M. (2005). Club boos op Charleroi, Het Nieuwsblad, September 29th (VUM).
- Ribeiro, C. & Urrutia, S. (2006). Scheduling the Brazilian soccer championship, in *PATAT' 06: Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling*, Brno, Czech Republic, pp. 481–483.
- Schreuder, J. (1992). Combinatorial aspects of construction of competition Dutch Professional Football Leagues, *Discrete Applied Mathematics* **35**(3): 301–312.